applied sciences

MDPI

Review

A Comprehensive Survey on Real-Time Image Super-Resolution for IoT and Delay-Sensitive Applications

M. J. Aashik Rasool , Shabir Ahmad, Sevara Mardieva, Sumaiya Akter and Taeg Keun Whangbo *

Department of IT Convergence Engineering, Gachon University, Sujeong-Gu, Seongnam-Si 461-701, Gyeonggi-do, Republic of Korea; aashikrasool@gachon.ac.kr (M.J.A.R.); shabir@gachon.ac.kr (S.A.); sevara1998@gachon.ac.kr (S.M.); sumaiya@gachon.ac.kr (S.A.)
* Correspondence: tkwhangbo@gachon.ac.kr

Abstract: In contemporary computer vision, deep learning-based real-time single image super-resolution approaches have gained significant attention for their ability to enhance the resolution of images in real time. These approaches are interconnected with various other computer vision domains, including image segmentation and object detection. Numerous surveys have summarized the state of the image SR domain. However, there is no survey that specifically addresses real-time single image SR on IoT devices. Therefore, in this study, we aim to explore strategies, identify the technical challenges, and outline the future directions of SR research, with a special emphasis on real-time super-resolution techniques. We begin with an overview of the core concepts related to real-time SR, recent challenges, and algorithm classification and delve into potential application scenarios that merit attention. Additionally, we explore the challenges and identify promising research areas related to real-time SR specifically related to IoT devices, highlighting potential advancements, limitations, and opportunities for future innovation in this rapidly evolving field.

Keywords: real-time image SR; image enhancement; real-time systems; single image super-resolution

Academic Editors: Chuan-Ming Liu and Chuan-Chi Lai

Received: 10 November 2024
Revised: 14 December 2024
Accepted: 26 December 2024
Published: 30 December 2024

Citation: Rasool, M.J.A.; Ahmad, S.; Mardieva, S.; Akter, S.; Whangbo, T.K. A Comprehensive Survey on Real-Time Image Super-Resolution for IoT and Delay-Sensitive Applications. *Appl. Sci.* **2025**, *15*, 274. https://doi.org/10.3390/app15010274

1. Introduction

Image super-resolution (SR) is a computer vision task focused on enhancing the resolution of images by transforming low-resolution (LR) visions into high-resolution(HR) ones [1–5]. To better understand and interpret the content of images, it is a significant challenge for researchers to derive high-quality images from LR ones [6–9]. SR technology has significant applications in various fields, particularly in medical imaging, where it enhances image quality for improved diagnosis and analysis. In medical imaging, SR techniques are employed to increase the resolution and clarity of various types of scans, including MRI, CT, and ultrasound [10,11]. This enhancement facilitates more accurate detection and diagnosis of medical conditions, potentially leading to better patient outcomes. Similarly, in the fields of surveillance and security, SR techniques play a pivotal role in improving video and image clarity, which is essential for enhancing situational awareness and ensuring public safety. By increasing the resolution of low-quality footage, these methods enable more effective monitoring in real-time scenarios, such as traffic surveillance, border security, and crowd management [12–14]. In addition, remote sensing and satellite imagery also benefit greatly from SR methods, which enhance image resolution for more accurate analysis and mapping. This improvement is critical for scientific fields such as astronomy, microscopy, and geology, where HR images lead to more precise observations and deeper insights [15–17].

On the other hand, with the widespread adoption of edge devices and real-time systems, the need for real-time SR approaches has increased. The primary objective of real-time image SR is to estimate an SR image from a given real-time LR image, ensuring that the generated SR image closely resembles the true HR image [18–21]. Currently, supervised learning-based SR networks are effective at increasing the resolution of real-time images. However, constructing real-time SR images presents significant challenges, such as high computational costs, which necessitate lightweight models [3,4,15,16,22]. These lightweight models have taken massive attention on the real-time SR domain [3,4,23–25]. In recent years, numerous researchers have published various reviews on single-image super-resolution (SISR) approaches that utilizes deep learning [1,2,26,27]. In ref. [28], the authors first categorized SR-based deep networks into nine distinct groups, including linear [29], residual, multi-branch [30], recursive, progressive, attention-based, and adversarial designs [31]. In addition, Garas et al. [32] categorized lightweight SR approaches into six groups: convolution-based, residual, dense, distillation, attention, and extremely lightweight models. Also, in [33], the authors reviewed the latest SISR methods and organized them into four main categories: RSISR based on degradation modeling, RSISR based on image pairs, RSISR based on domain translation, and RSISR based on self-learning. Furthermore, in [14], the authors provide an in-depth discussion of blind image SR by categorizing existing methods into three classes according to the degradation modeling and the data employed in solving the SR model [34].

However, these reviews have generally focused on frameworks that may or may not be deployable as lightweight models [35], without considering the specific applications for which the models are used in real time [36]. Additionally, while previous surveys have examined existing SR methods from various perspectives [37], to the best of our knowledge, none have specifically focused on real-time SR for delay-sensitive applications [38]. Considering this, we aim to review and summarize advancements in the field of real-time SR for delay-sensitive applications.

Contributions and Scope

This survey presents an overview of the latest advancements in real-time SR techniques. The goal is to offer foundational knowledge and set the stage for future developments in real-time SR tasks. It aims to identify key variations and parameters that could be considered for future innovative approaches. While there are many valuable surveys covering various aspects of SR, there is a notable lack of comprehensive sources specifically focused on real-time SR and emerging methodologies in this area.This study aims to address this gap.

2. Background

2.1. Overview of Super-Resolution

SR is a method used to increase the resolution of digital photographs and videos in order to improve their quality [39]. This technique includes creating HR images from LR inputs, either from collected frames or in real-time [40]. It was developed from research conducted at AT&T Bell Laboratories in the 1990s. "Super-resolution" describes the technique of creating digital frames with HR to improve the spatial detail of the source image.

In mathematical terms, if we denote an LR image as LR_x, the degradation process can be modeled using a function D. The corresponding HR image is represented as H_x, and noise is accounted for by the variable σ. This relationship can be expressed as Equation (1). To achieve HR outputs, various methods are employed, including reconstruction-based methods, interpolation-based methods, learning-based methods, example-based methods, and hybrid methods.

$$LR_x = D(H_x; \sigma) \quad (1)$$

Several well-known methods are used to construct F outputs, including reconstruction-based, interpolation-based, learning-based, example-based, and hybrid-based methods. Figure 1 shows the taxonomy of the SR techniques and popular examples of each branch.

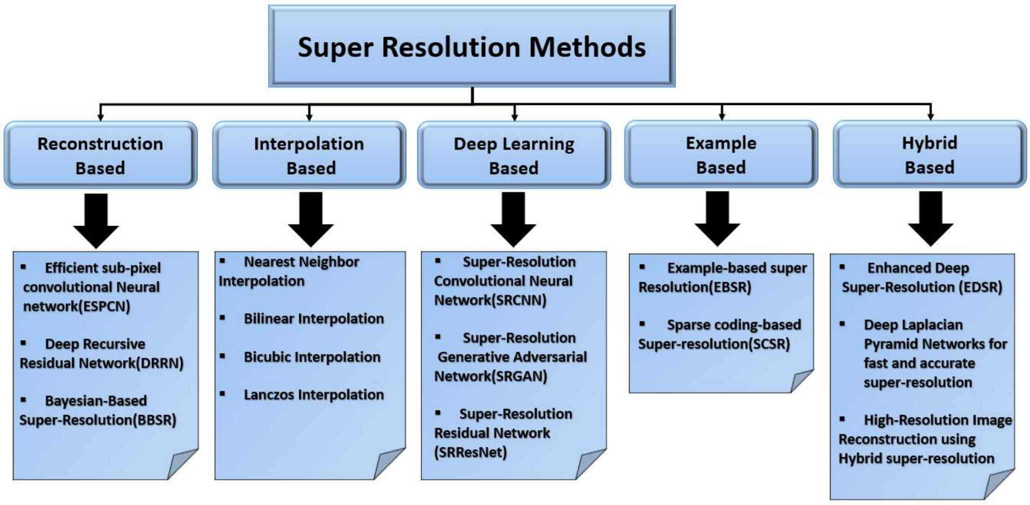

Figure 1. Taxonomy of the super-resolution techniques.

Reconstruction-Based Super-Resolution Method

Reconstruction-based SR methods are a subclass of SR that minimizes the difference between LR images and estimates the HR images [5]. This method uses an iterative approach that requires prior knowledge regarding the image in order to recover the HR frequencies that were lost during the process of LR image acquisition [41]. These methods utilize mathematical models that consist of a forward model and regulation term. In forwarding model, it describes the image formation process. This forward model consists of key factors such as the point spread of the system imaging, resolution of sensors, etc. On the other hand, a regulatory term is a mathematical function that is used in conjunction with a forward model to improve the smoothness of the HR picture. The optimization problem is defined by minimizing the dissimilarity between the LR images and the predicted HR image. This is subject to a regularization term. There are several well-known approaches introduced under reconstruction-based SR methods such as the Efficient Sub-Pixel Convolutional Neural Network (ESPCN) [42], Deep Recursive Residual Network (DRRN) [24], and Bayesian Image Super-Resolution (BSR) [43].

The ESPCN architecture was introduced to upsample an LR image to an HR image by utilizing sub-pixel convolutional layers [42]. In this context, convolutional layers can learn to rearrange LR feature maps to generate HR images directly. This process reduces computational complexity. Apart from the ESPCN method, the DRRN is one of the popular methods that utilizes the Reconstruction-based SR method. In this approach, they constructed a recursive network that consists of several residual blocks. Every residual block consists of batch normalization and rectified linear unit (ReLU) activation layers. In 2001, the authors of [43] introduced the Bayesian framework for image SR, which can utilize prior distribution over HR images to regulate the solution. In this approach, they use a Laplacian distribution as the prior, which is characterized by a greater degree of smoothness in the HR image. In this case, the network applies the input image using a two-step process to generate the HR output. As a first step, it upsamples the LR input by utilizing the linear interpolation method. The second stage uses a Bayesian framework to regulate the solution utilizing a prior distribution across HR photos.

2.2. Interpolation-Based SR Method

Interpolation-based techniques of SR include nearest neighbor interpolation, bilinear interpolation, bicubic interpolation, and Lanczos interpolation [44]. These methods generate the initial HR estimate of the output image by utilizing interpolation methods. In this step, the interpolation stage entails employing upsampling techniques and generating the HR output.

In the nearest neighbor interpolation method, the concept underlying such algorithms is that comparable image are near in terms of distance measurement [45]. The nearest neighbor interpolation approach can be represented mathematically as follows in Equation (2):

Given a collection of datasets $(L_1, M_1), (L_2, M_2), ..., (L_n, M_n)$ and a position L, the estimated value M at L is

$$M = M_i \text{ if } L_i \leq L < L_{i+1}, \text{ for } i = 1, 2, \ldots, n-1, M = M_n; \text{ if } L \geq L_n \tag{2}$$

In this case, M_i indicates the M value of the data point closest to L, while L_n represents the data set's greatest L value [46].

Apart from the nearest neighbor interpolation method, bilinear interpolation is one of the well-known techniques for producing HR output [47]. The bilinear interpolation approach is accustomed to generating an HR output straightforwardly and efficiently. A bilinear interpolation is one of the comprehensive techniques for two-dimensional linear interpolation. The value at one place inside the rectangular grid is calculated using a weighted average of the values at the four closest grid points using bilinear interpolation. The distance between the interpolated location and the surrounding grid point positions determines the weights applied to each grid point.

On the other hand, bicubic interpolation is a method of interpolating data points on a two-dimensional regular grid that is an extension of cubic spline interpolation. This approach is often used in SR tasks due to computational efficiency and comparably better performance than other interpolation methods. Apart from these methods, Lanczos interpolation is a well-known method for upscaling the LR images [48]. The intuition behind the method is the Lanczos Kernel, which is a windowed sinC function with superior frequency domain features compared to conventional interpolation kernels. In particular, the Lanczos method was utilized for sharpness preservation and is customizable.

2.3. Deep Learning-Based Super-Resolution Method

The intuition behind the deep learning-based SR is utilizing an advanced neural network algorithm and enhancing the LR digital frame's resolution. In this type of approach, the deep neural network is trained on a large dataset of HR image and its corresponding LR images. This approach allows for mapping between LR image to HR image.

The initial deep learning-based algorithm was introduced in 2014 [49] as image SR using deep convolutional networks (SRCNNs). In this approach, the authors designed the network in three consistent layers: the first layer extracts the features from the LR image, the second layer does non-linear mapping, and the last layer constructs the HR counterparts. The relationship is represented in Figure 2.

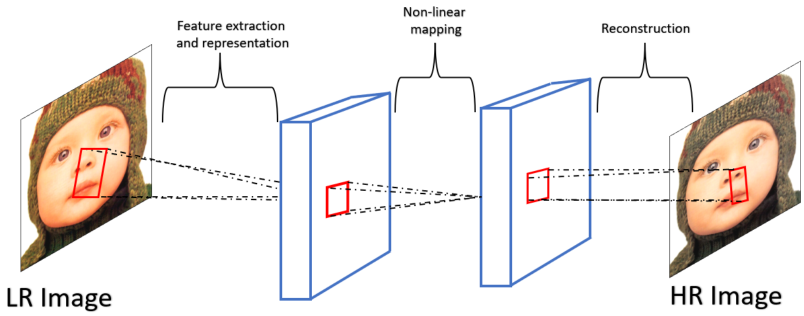

Figure 2. Architecture of SRCNNs.

Apart from SRCNNs, the first Generative Adversarial Network (GAN)-based approach was introduced by Ledig et al. [50] in 2016. In this method, authors designed a generator that upsamples the LR images after the 16 residual convolutions. Here, each residual block consists of a convolutional layer, batch normalization, PReLU activation function, convolutional layer, batch normalization, and element-wise summation layer, respectively. Here, a generator produces an HR image, and then the discriminator classifies whether the generator's produced image is real or fake. In the discriminator, authors followed eight residual blocks, where each residual block consists of convolution, batch normalization, and Leaky ReLU followed by each other.

2.3.1. Exampled-Based SR Method

The example-based super-resolution (EBSR) method utilizes a database of HR sample images to increase LR inputs [51]. This method follows a data-driven approach that utilizes the content information in the sample images to construct HR images. The EBSR contains six sub-approaches such as dataset collection, registration, patch extraction, patch similarity measure, patch fusion, and post-processing. In the dataset collection part, create a sample collection of HR images that resembles the images that are intended to enhance. In the registration part, the LR input image is aligned with the reference dataset's instances. This alignment is required to generate correspondence between the LR and HR images. In the patch extraction part, the LR input image and the HR reference image are separated into tiny overlapping images. Based on a similarity metric, such as pixel intensity or structural similarity, each LR patch is matched to similar HR images in the provided dataset. Apart from this image similarity measure working for each LR patch, a similarity metric is utilized to locate the most comparable HR patches from the reference dataset. This metric might be based on pixel intensity, gradient information, texture qualities, or other image properties. On the other hand, patch fusion works as a re-constructor in which the HR version of the image and HR patches from the reference dataset are fused together. Averaging, weighted mixing, and patch-based synthesis are some of the fusion techniques that may be applied. As a last part, post-processing works as the resultant HR digital frames to improve visual quality, eliminate artifacts, or polish features. Denoising, sharpening, and adaptive filtering are examples of such approaches.

Apart from EBSR, sparse coding-based super-resolution (SCSR) is one of the well-known techniques used to enhance LR digital frames to HR digital frames. The main intuition behind the network is sparse coding, wherein natural digital frames can be sparsely displayed utilizing basis functions. SCSR methods target search sparse representation of LR digital frames in order to the dictionary, afterward utilizing this representation to construct HR digital frames. The above process involves producing solutions to optimization issues by finding the sparsest representation.

2.3.2. Hybrid-Based Super-Resolution Method

Hybrid-based SR techniques are referred to as a class of methods that utilize multiple components or approaches to construct HR digital frames. Figure 3 demonstrates the taxonomy of hybrid-based SR methods.

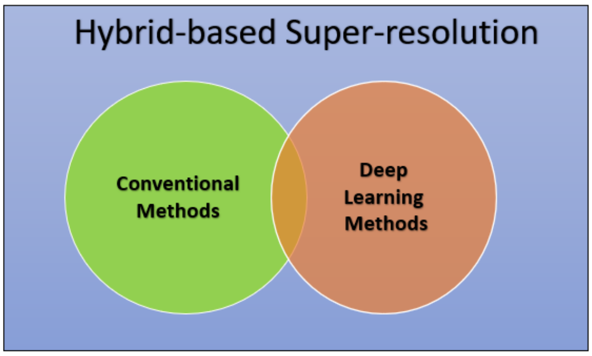

Figure 3. Taxonomy of hybrid-based SR techniques.

There are two components of hybrid-based SR techniques, including conventional methods and deep learning-based methods. Conventional methods contain the traditional signal processing methods, for instance, interpolation, edge enhancement, and regularization methods. Secondly, deep learning-based techniques encompass the utilization of the use of deep neural networks, such as convolutional neural networks (CNNs), to learn about the link between LR and HR images.

Enhanced Deep Super-Resolution (EDSR) [52] is an example of a hybrid-based SR technique that utilizes deep learning-based approaches. EDSR uses a deep neural network architecture utilizing CNNs to study the LR and HR images. EDSR consists of multiple convolutional layers including several residual blocks. This process helps to smooth the flow of gradients and aids in the convergence of the network. Apart from EDSR, deep Laplacian pyramid networks for fast and accurate SR (LapSRNs) [53] represent a well-known hybrid-based SR method. Traditional methods use iterative optimization techniques or sophisticated models, which can be computationally and time costly. The authors try to overcome these constraints by utilizing deep learning and a multi-scale architecture. LapSRNs are trained to utilize pairings of LR and HR images in a supervised way. The network studies how to reduce the reconstruction error between the expected HR image and the ground truth during training. Typically, the loss function contains both pixel-wise and perceptual changes. This approach offers a balance of precision and speed, making it suited for a wide range of real-world applications.

2.4. Datasets and Assessment Metrics

2.4.1. Datasets

Prior to 2017, only a few datasets, such as the T91 image dataset [54] and the Berkeley Segmentation Dataset (BSD) [55], were commonly used in SR tasks. Real-time SR models were primarily trained using these datasets. However, after the introduction of the DIV2K dataset [56] in the NTIRE 2017 competition [55], it became the main dataset for training most lightweight SR algorithms. Table 1 shows the most recent datasets used to train and evaluate SR models along with the explanation of the datasets.

Table 1. Explanation of each dataset widely used in real-time super-resolution.

Name of Dataset	Explanation of Dataset
DIV2K	A popular dataset for SR tasks consisting of 1000 HR RGB images. It includes 800 training, 100 validation, and 100 test images. The dataset is designed for benchmarking and evaluating image SR algorithms, as well as other image restoration tasks. All images in the dataset are LR and are used as ground truth for generating lower-resolution images, which are then used to train and test SR algorithms.This dataset is essential for developing and testing new methods in image processing and computer vision due to its diverse range of high-quality images and standardized resolution.
Set5	A dataset composed of five HR images: baby, bird, butterfly, head, and woman.This dataset was designed for benchmarking and testing SR algorithms. It provides a small, standardized set of images to evaluate the performance of image enhancement techniques. It is commonly used for testing SR methods. This dataset is crucial for establishing benchmarks in SR research, providing a standardized set of images for consistent and comparable evaluation of different techniques.
Set14	Similar to Set5, Set14 includes 14 different types of HR images.The images cover a broad range of categories, including landscapes, urban scenes, and various types of textures. The images are usually available in common formats such as JPEG or PNG. The images in Set14 are HR and serve as ground truth for creating lower-resolution versions.
BSD100	This collection contains 100 HR photos from Berkeley's Segmentation collection. The collection comprises natural photos with a wide range of items, textures, and situations. Researchers and developers frequently use BSD100 to assess the performance of their image processing models by running their algorithms on the dataset and comparing the results to ground truth or other models using common metrics such as the PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index).
Urban100	Urban100, which consists of 100 HR urban scene images, is noted for its richness and complexity, covering a wide range of sizes and architectural styles. Urban100 is mostly used to evaluate the performance of image SR techniques. Researchers utilize the Urban100 dataset to assess and compare the efficacy of various SR methods.
CelebA	A large-scale dataset of celebrity faces often used for training models that perform SR on human faces.These attributes include characteristics like gender, age, hair color, the presence of eyeglasses, facial expressions, and more. It has been used to train and evaluate models for tasks such as Attribute Prediction,Generative Models, and Face Recognition.
Flickr2K	A high-quality image dataset containing 2650 HR images (2K resolution). It can be used for various image tasks, including SR. Researchers use it to train models and evaluate performance.
COCO dataset	The COCO dataset contains a wide variety of images and is commonly used for tasks such as object detection, segmentation, and SR. COCO is freely available and widely used in both academic and commercial research. The dataset is continuously updated, with the most recent versions including even more annotations and categories.
ImageNet	ImageNet is a substantial visual database used in the field of visual object recognition and classification research. It contains over 14 million annotated images across more than 20,000 categories. The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) has been conducted using the ImageNet dataset since 2010, allowing participants to compete in tasks related to object and scene classification, as well as detection.

2.4.2. Assessment Metrics for Super-Resolved Images

In the quality assessment of SR images, there are two primary approaches: one is subjective evaluation based on human perception, and another one is object evaluation using quality metrics. Subjective evaluation aligns more closely with practical needs but has drawbacks, such as being influenced by personal preference and being costly and difficult to automate. On the other hand, objective evaluation is more convenient and can be automated, although different metrics may yield inconsistent result compared to each other and subjective evaluation. These reports outline key metrics used to objectively evaluate the quality of SR images [57], SSIM [58], LPIPS [59], and NIQE [60]. In this context, let $\mathbf{X} \in \mathbb{R}^{H\times W\times C}$ and $\hat{\mathbf{X}} \in \mathbb{R}^{H\times W\times C}$ denote the ground truth image and the super-resolved image, respectively. H, W, and C denote height, width, and the number of components, respectively. The PSNR is the most commonly used statistic to assess the quality of reconstructed images, especially in SR and image compression. It calculates how similar the original and the compressed or rebuilt image are to each other. (e.g., SR, denoising, deblocking, and deblurring). Given $\hat{\mathbf{X}}$ and $\mathbf{X}$, the PSNR is defined as

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{L^2}{\mathrm{MSE}}\right) \tag{3}$$

where

$$\mathrm{MSE} = \frac{1}{HWC}\|\mathbf{X} - \hat{\mathbf{X}}\|_2^2 \tag{4}$$

Equation (4) denotes the mean squared error (MSE) between $\hat{\mathbf{X}}$ and $\mathbf{X}$, and L represents the maximum pixel value (i.e., 255 for 8-bit images). It can be seen from Equation (3) that the PSNR is more concerned with the proximity between corresponding pixels in $\hat{\mathbf{X}}$ and $\mathbf{X}$, which results in low consistency with perceptual quality in some cases.

The SSIM [57] is a full-reference objective quality metric titled the Structural Similarity Index (SSIM). The authors in [57] used it to compare the structural similarity of two images. By analyzing the contrast, brightness, and structure of the original image $\mathbf{X}$ against the rebuilt image, it measures the quality of the image.

$$\mathrm{SSIM} = \left[l(\mathbf{X},\hat{\mathbf{X}})\right]^{\alpha}\left[c(\mathbf{X},\hat{\mathbf{X}})\right]^{\beta}\left[s(\mathbf{X},\hat{\mathbf{X}})\right]^{\gamma} \tag{5}$$

where

$$l(\mathbf{X},\hat{\mathbf{X}}) = \frac{2\mu_{\mathbf{X}}\mu_{\hat{\mathbf{X}}} + C_1}{\mu_{\mathbf{X}}^2 + \mu_{\hat{\mathbf{X}}}^2 + C_1}, \quad c(\mathbf{X},\hat{\mathbf{X}}) = \frac{2\sigma_{\mathbf{X}}\sigma_{\hat{\mathbf{X}}} + C_2}{\sigma_{\mathbf{X}}^2 + \sigma_{\hat{\mathbf{X}}}^2 + C_2}, \quad \text{and} \quad s(\mathbf{X},\hat{\mathbf{X}}) = \frac{\sigma_{\mathbf{X}\hat{\mathbf{X}}} + C_3}{\sigma_{\mathbf{X}}\sigma_{\hat{\mathbf{X}}} + C_3}.$$

C α, β, and γ are weighting parameters; $\mu_{\mathbf{X}}$ and $\sigma_{\mathbf{X}}$ denote the mean and standard deviation of $\mathbf{X}$, respectively. Similarly, $\mu_{\hat{\mathbf{X}}}$ and $\sigma_{\hat{\mathbf{X}}}$ denote the mean and standard deviation of $\hat{\mathbf{X}}$. $\sigma_{\mathbf{X}\hat{\mathbf{X}}}$ is the covariance between $\hat{\mathbf{X}}$ and $\mathbf{X}$. C_1, C_2, and C_3 are constants.

Furthermore, Equation (5) can be simplified when $\alpha = \beta = \gamma = 1$ and $C_3 = \frac{C_2}{2}$ as

$$\mathrm{SSIM} = \frac{(2\mu_{\mathbf{X}}\mu_{\hat{\mathbf{X}}} + C_1)(2\sigma_{\mathbf{X}\hat{\mathbf{X}}} + C_2)}{(\mu_{\mathbf{X}}^2 + \mu_{\hat{\mathbf{X}}}^2 + C_1)(\sigma_{\mathbf{X}}^2 + \sigma_{\hat{\mathbf{X}}}^2 + C_2)}. \tag{6}$$

In general, the SSIM [57] is thought to more accurately represent visual quality than the PSNR. For evaluating the quality of a restored image, the PSNR and SSIM [57] are typically utilized in combination when the corresponding ground truth image is accessible.

A full-reference metric titled the Information Fidelity Criterion (IFC) [58] evaluates image quality using data from natural scenes. Using models that describe natural situations and their distortions, the mutual information between the test image and the reference image is measured to assess the visual quality. Research indicates that the IFC [58] is a useful metric for evaluating super-resolved image quality [61]. LPIPS [61]: Learned Perceptual Image Patch Similarity, or LPIPS, is a complicated metric that evaluates perceptual similarity to assess image quality. LPIPS evaluates image quality by calculating the L2 distance between deep feature representations of the reference and test images, which is in contrast to conventional approaches that compare images pixel-by-pixel [61]. Using characteristics taken from neural networks that were previously trained, this method closely resembles human vision and offers a more accurate depiction of perceived image quality.

The Natural Picture Quality Evaluator, often known as NIQE [60], is a completely blind picture quality evaluation metric that does not depend on particular aberrations or human opinion. It evaluates "quality-aware" features taken from images utilizing a multivariate Gaussian (MVG) model. These features include parameters from asymmetric and generalized Gaussian distributions. The MVG model fitted to the assessed image and the MVG model fitted to the natural image distribution are compared, and the difference between the two is used to determine the quality of the image.

3. Real-Time Systems and Real-Time SR

3.1. Real-Time Systems

The intuition behind the real-time system operates in accordance with real-world time constraints and requirements such as that the response time should meet some specified restriction of time or systems that adhere to certain deadlines. When considering timing constraints, different types of real-time systems can be identified such as hard real-time systems, soft real-time systems, and firm real-time systems. A hard real-time system is a system characterized by severe time requirements when missing deadlines might have devastating effects. Tardiness, or the delay in performing tasks relative to their deadlines, reduces the utilization of system outcomes. Examples of systems in this category include automated driving systems and avionics systems. Apart from the hard real-time system, soft real-time systems are a well-known method that utilizes real-time systems. This sort of system may occasionally fail to meet its deadline with an acceptable degree of probability. There are no dire penalties for missing the deadline. Examples of systems in this category include streaming media systems, interactive video games, and web servers. Real-time SR also falls into this category. Apart from soft real-time systems, firm real-time systems are a popular method that utilizes real-time systems. The type of systems fall in between hard and soft real time. Missing a deadline is bearable in solid real-time systems, but the utility of the output declines with delay. Systems that fall into this category include telecommunications switching systems, real-time financial trading systems, and manufacturing control systems. Figure 4 demonstrates the relationship between the real-time systems.

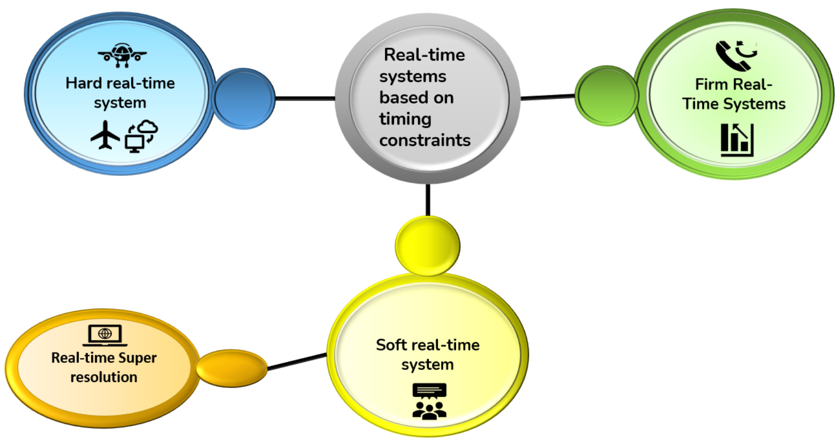

Figure 4. Relationship between the real-time systems based on timing constraints.

Real-time SR is classified within the category of soft real-time systems. Real-time SR algorithms have the capacity to provide improved digital frames within the required period. These strategies can handle minor delays without causing substantial harm. This compromise enables satisfying time constraints while maintaining image resolution.

3.2. Delay-Sensitive Applications

Applications that rely on timely data delivery and processing are referred to as delay-sensitive applications. Due to their sensitivity to latency, even minor delays can significantly impact their usability or performance. Although the significance of meeting deadlines may vary, these applications generally require data handling or responses within strict time constraints to ensure proper operation. This typically involves assessing the performance of the operating system, firmware, network, and hardware—both individually and collectively—for the entire system [62]. Low latency is beneficial for latency-sensitive applications, as it signifies a minimal delay between the initiation of an action and its result. High latency is undesirable because it can slow down data transmission and may indicate issues such as packet loss or droppage depending on the type of data. Even if the average latency seems low, a network with excessive jitter is unreliable. Therefore, latency should also be consistent.

3.3. Real-Time Single Image Super-Resolution Methods

Enhancing the spatial resolution of images to improve their clarity and detail is the primary aim of real-time SR. Several strategies are used to achieve this. Distillation makes real-time processing more feasible by training a smaller, more efficient model to replicate the performance of a larger, more complex model. Re-parameterization improves effectiveness and performance by modifying the model's parameters, often by transforming or simplifying the network architecture. Model-based techniques leverage prior knowledge or assumptions about the data, sometimes incorporating mathematical models or constraints, to guide the resolution enhancement process. Quantization reduces the precision of the model's parameters, which accelerates processing times by using less memory and computational power. Lastly, optimization techniques utilize various algorithms to fine-tune the model's performance, balancing resource usage, accuracy, and speed. Figure 5 illustrates the taxonomy of real-time SR methods.

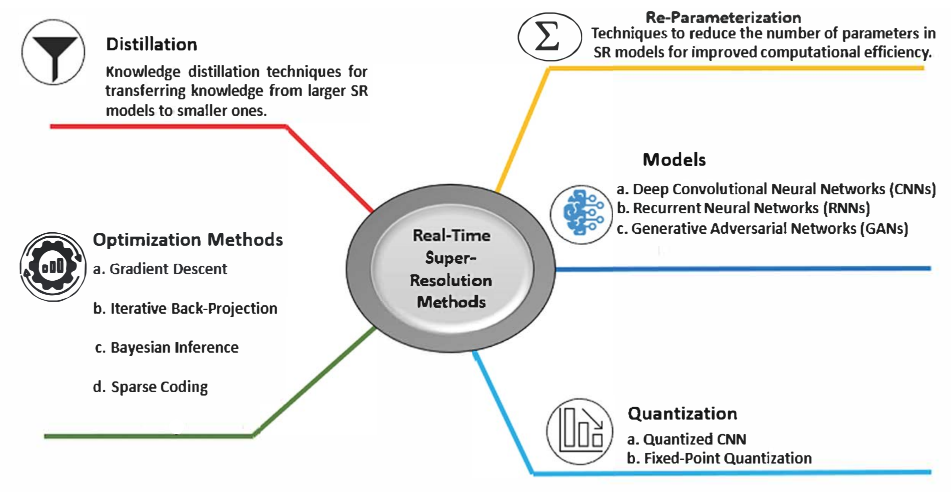

Figure 5. Taxonomy of real-time super-resolution methods.

3.3.1. Distillation-Based Real-Time Super-Resolution Methods

Real-time SR based on distillation is a strategy that utilizes knowledge distillation to construct HR SR outcomes in real-time circumstances. These types of models follow the approach of student and teacher models. The intuition behind these models, training a faster and tiny model that behaves like a student and mimics the attributes and performance of larger models, is to create more accurate SR tasks. Figure 6 displays how knowledge distillation works in a tiny student model to replicate a big teacher model and harness the teacher's knowledge to achieve comparable or superior accuracy. Here, larger models act as teachers. Distillation-based SR technologies frequently use lightweight network designs and optimization approaches to attain real-time performance. Some solutions, for example, make use of network topologies with a limited number of parameters, such as MobileNet or ESPCN.

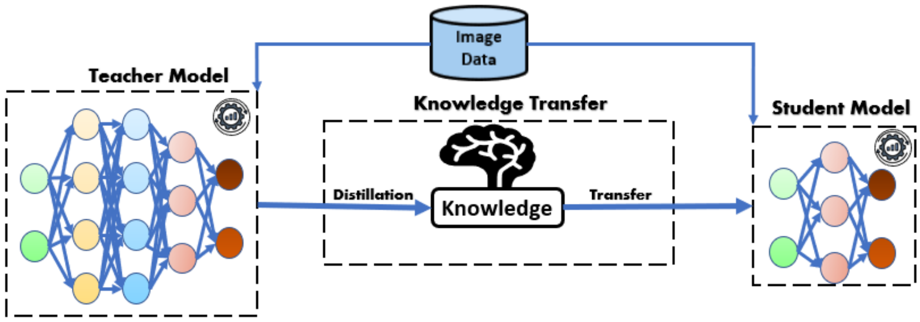

Figure 6. The framework for knowledge distillation work through the teacher–student model.

The authors of [63] presented the first model based on information distillation, known as the Information Distillation Network (IDN). The IDN architecture is made up of three major components: feature extraction blocks, stacked information distillation blocks (DBs), and reconstruction blocks. Within the DB, an enhancement unit and a compression unit collaborate to efficiently extract both local long- and short-path characteristics. Specifically, the enhancement unit combines two types of characteristics, whereas the compression unit distills more usable information through consecutive blocks. The IDN model has a noteworthy benefit in that it uses very few layer filters and group convolution. However, the IDN still has a significant weight, emphasizing the need for more efficient models while deploying in real-time systems.

In [64], the authors proposed a feature-domain adaptive method utilizing contrastive distillation for lightweight SISR. They addressed the limits of previous feature distillation approaches that use Euclidean distance-based loss and developed a feature-domain contrastive loss. This novel technique allows student networks to learn more detailed information from the teacher's representation in the feature domain and be efficient while deploying on real-time systems. As a solution for IDN's obstacles, the authors of [65] presented the Information Multi-Distillation Network (IMDN) for SISR, which uses cascaded Information Multi-Distillation Blocks (IMDBs). This architecture allows for the sequential extraction of hierarchical features using a distillation module. These traits are subsequently integrated by a fusion module, which prioritizes them according to their importance. A Contrast-Aware Channel Attention (CCA) layer is used for task assessment. Furthermore, an Adaptive Cropping Strategy (ACS) is created to handle actual photos of various sizes using the same trained model. While the IMDN overcame some of the difficulties associated with the IDN's high weights, there is still significant room for improvement in the IMDN paradigm.

In [15], an innovative approach named the Residual Feature Distillation Network (RFDN) is described, which uses a channel-splitting operation to build a Feature Distillation Connection Block (FDCB). This block is an upgraded version of the previous IMDN [65]. The RFDN focuses on learning discriminative feature representations by combining numerous feature distillations with a shallow residual block (SRB). The SRB, which consists of a convolutional layer, a skip connection, and an activation unit, uses residual learning to improve performance. This addition marginally enhances computation without adding any additional parameters compared to the original IMDN block. Additionally, the model incorporates an Enhanced Spatial Attention (ESA) mechanism to further improve performance. As a result, the RFDN network is lighter in weight and runs faster than the IMDN. In addition, Haoran Yang et al. [66] proposed a information distillation network based on feature similarity ranking for lightweight image SR. They presented a feature similarity ranking system that organizes extracted feature channels according to their degree of redundancy. The ranked information distillation block (RIDB) is then utilized to extract features with a high redundancy level. Furthermore, since the feature channels have already been sorted in the channel dimension, a gated non-local attention module (GNLM) is used to capture long-distance feature correlations in the spatial dimension. The GNLM considerably decreases computing costs by employing a partitioned nonlocal module and a gate unit to assess the significance of each module, making it more appropriate for lightweight real-time SR. Apart from this, Min Yoon et al. [67] introduced a framework utilizing the teacher assistant knowledge distillation (KD) approach for SR. When the HR image is put into the teacher model to outperform the student, a considerable performance gap opens up due to the teacher's better skills. As a result, the teacher's information becomes extremely complicated and difficult to impart, thereby reducing the efficacy of the KD. To address this, they used a teacher–student (TS) network to increase knowledge transfer between the teacher and the students. Furthermore, the distribution of compact features (CFs), which are the inputs to the teacher's decoder, was discretized to match the student's input distribution, thereby improving the KD process. By deploying this approach, they significantly reduced the complixity and made it suitable for deployment in real time.

Discussion and Limitations

Feature distillation models have improved gradually, starting with the IDN, which simply utilized basic information distillation. This transitioned to the IMDN, which included various distillation processes. Building on the IMDN, residual learning was included, resulting in the RFDN. Subsequent feature distillation models improved the SRB established in RFDN. These KD models are based on the idea of taking existing SR models and decreasing their weight and computational needs utilizing the TS structure. Furthermore, the proper initialization of student models is critical for training these models. Furthermore, ref. [66] used novel feature similarity ranking to increase task performance. While distillation-based approaches that use channel splitting for feature distillation have shown good results, they are limited by the convolution kernel, which restricts their ability to capture long-range relationships. To solve this issue, attention-based approaches have been developed, which enable the identification of long-range relationships using blocks capable of extracting non-local information.

3.3.2. Optimization-Based Real-Time Super-Resolution Methods

Optimization-based real-time SR methods are one of the popular subclasses of real-time SR techniques. These approaches utilize the SR problem as an optimization task, intending to locate the HR images from LR inputs by minimizing a cost function that balances fidelity to the observed data with smoothness or regularization terms. There are five types of optimization techniques utilized in real-time SR tasks, such as gradient descent, iterative backprojection, Bayesian inference, sparse coding, and convex optimization.

Gradient Descent-Based Optimization Methods

Gradient descent-based optimization methods improve the performance of real-time SR. There are several approaches conducted to optimize the real-time SR. In [68], the authors tackled the perception–distortion trade-off in SR as a multi-objective optimization issue. They developed a novel optimizer by combining a gradient-free evolutionary algorithm (EA) with the gradient-based Adam optimizer. The EA facilitates the exploration of alternative optimization tactics, whereas Adam improves their accuracy. This method provides a set of ideal models that cater to various balancing points between perception and distortion. The authors then created a fusion network to merge these models into one more powerful model, establishing a better balance between perception and distortion, which is well-suited for lightweight deployment. Apart from this, [69] investigated image restoration and synthesis using gradient descent algorithms. They examined at both linear and non-linear restoration challenges, as well as how to incorporate a known degradation model into a gradient-based restoration method.To investigate gradient-based image restoration, researchers employed a bicubic SR. It tested real bicubic degradation against CNN and linear approximations, utilizing MSE as the loss function. The Residual-in-Residual Dense Block (RRDB) architecture handled the restoration using five gradient descent steps to recover images from degraded inputs.

Bayesian Inference Methods

Bayesian inference methods are rarely used in real-time SR tasks due to their computational complexity. Gaussian processes (GPs) are used to represent the data frame space at HR and give a probabilistic foundation for SR. GPs are non-parametric models that are flexible enough to represent the uncertainty in the SR problem. GPs may be used to estimate the HR digital frame and offer uncertainty estimates associated with the reconstructed image by considering the SR process as a Bayesian inference issue.

Weixin Li et al. [70] presented a low-dimensional sparse Bayesian learning with a Doppler compensation (LDSBL-DC) approach for improving azimuth resolution in aerial forward-looking imaging while reducing computing complexity. The solution addresses the issue of Doppler centroid variation induced by pitching angle changes through generating a Doppler compensation matrix that removes this spatial variance. This permits the Doppler convolution matrix to be created only once. In addition, they present a low-dimensional projection model that use singular-value decomposition to compress high-dimensional echo data into a lower-dimensional form. By merging Doppler correction with this projection model, a novel imaging technique is developed, with imaging parameters estimated via sparse Bayesian learning (SBL). Matrix transformations are used to further minimize computing complexity when estimating target scattering coefficients. The method's low computing complexity, obtained by Doppler correction and low-dimensional projection, allows for rapid processing of forward-looking imaging data. This makes them easier to integrate into real-time systems, where fast and precise image processing is critical. In addition, ref. [71] introduced a Bayesian deep learning technique for assessing and comprehending uncertainty in SR-guided wave array imaging. It uses a Monte Carlo (MC) dropout approach within multi-scale deep learning models for approximate Bayesian inference, which helps to quantify uncertainties in imaging subwavelength defects. This technique distinguishes between aleatoric uncertainty, which stems from the data, and epistemic uncertainty, which is linked with the model itself.

Sparse Coding Methods

The method of expressing signals as a linear combination of a few basis functions, with most coefficients close to zero, is referred to as sparse coding. In the real-time SR domain, sparse coding methods can be utilized in two steps such as dictionary learning and sparse reconstruction. Dictionary learning entails locating an over-complete set of basis functions, referred to as a dictionary, that can sparsely represent HR picture patches. A training set of HR photos is used to train the dictionary. Various dictionary learning algorithms, such as the K-SVD algorithm and the online dictionary learning method, can be employed. The learned dictionary captures the underlying patterns and structures in HR space, enabling sparse representation to be economical. Secondly, it follows sparse reconstruction. The sparse reconstruction stage uses the learned dictionary to identify the sparse representation of the patches given an LR digital frame. This can be accomplished by solving a sparsity-promoting optimization problem, such as an L1-norm minimization problem. The sparse reconstruction approach improves the representation of high-frequency information that is lacking in the LR input by increasing sparsity.

Sparse arrays offer better resolution and degrees of freedom compared to traditional uniform linear arrays (ULAs). Pulak Sarangi [72] refutes the claim that sparse arrays need many temporal measurements for accurate parameter estimation. Through a non-asymptotic analysis of the Coarray ESPRIT algorithm, his work shows that coarray manifold scaling reduces estimation errors even with fewer snapshots. For nested arrays with fewer sources than sensors, estimation errors can be minimized with a logarithmic number of snapshots relative to sensors. The study also confirms that sparse arrays can achieve source separation with a minimum distance of $\Omega(1/P^2)$ versus $\Omega(1/P)$ for ULAs, and this highlights their superior noise resilience and sample complexity dependence on the Signal-to-Noise Ratio (SNR) and source power. This approach will guide to efficient deplyment in real-time SR.

In [25], the authors investigated at how incorporating sparsity into image SR might improve the efficiency of SR networks during inference. They proposed the Sparse Mask SR (SMSR) network, which is meant to generate sparse masks and decrease duplicate processing. Specifically, the network employs spatial masks to identify "important" parts in the image, whereas channel masks detect redundant channels in less-critical locations. This approach allows the network to choose to avoid duplicates operations, resulting in identical performance but with improved efficiency.

Discussion and Limitations

Optimization-based real-time SR approaches are successful, but they have several drawbacks. Gradient descent approaches, despite improving the perception–distortion trade-off, frequently use predetermined degradation models and are hyperparameter-sensitive, restricting their versatility. Bayesian inference, while informative, is computationally demanding and unsuitable for real-time applications. Sparse coding approaches improve high-frequency detail but are strongly reliant on the quality of the learned lexicon. Despite advances such as SMSR to enhance efficiency, obtaining real-time performance without losing picture quality remains difficult, necessitating more adaptable and computationally efficient techniques.

3.3.3. Re-Parameterization in Real-Time SR

Re-parameterization is a technique used in real-time SR to simplify the SR issue and make it easier to optimize or estimate. The computational complexity of the issue can be lowered by re-parameterizing it, allowing for real-time performance. There are two types of re-parameterization methods available in real-time SR tasks: scale-space re-parameterization and sub-pixel re-parameterization. The SR issue is altered in scale-space re-parameterization by creating intermediary scales between the LR and HR digital frames. Rather than instantly upsampling the LR input to the intended HR, the re-parameterization process progressively upsamples the input to numerous intermediate scales. By dividing the SR work across different scales, this technique minimizes computing complexity and allows for real-time processing.

Secondly, sub-pixel re-parameterization is frequently utilized to address the constraints of pixel-based SR approaches. Instead of executing SR at the pixel level, this approach attempts to estimate sub-pixel shifts or features in the LR input. The re-parameterized issue becomes easier to solve by estimating sub-pixel shifts, and the subsequent upsampling procedure can provide more accurate and visually appealing HR images.

In [73], the authors present an edge-priority-extraction network established with their indicated edge-priority blocks (EPB). The EPB improves network representation by combining numerous branches that include edge information. Additionally, the EPB may be re-parameterized for more efficient inference. To make better use of edge information, the study offers a mixed-priority filter that extracts edge information based on horizontal and vertical priorities, boosting network performance. These filters can adapt and capture edge information using multi-directional derivatives. With this approach, they succeeded in reducing the model's complexity, making it suitable for real-time deployment. In addition, ref. [74] presents a re-parameterized dynamic unit (RDU) as a plug-in component with the goal of improving the performance–inference cost trade-off. During the training phase, the RDU learns to combine numerous re-parameterizable blocks by assessing various input data, hence improving layer-level representation. In the inference step, the RDU is transformed into simple dynamic convolutions capable of capturing both dynamic and static feature maps. These RDUs are then combined to build an information distillation block that enables hierarchical refining and selective fusing of spatial context information. Furthermore, the authors suggest a dynamic distillation fusion (DDF) module that allows for dynamic signal aggregation and communication between hierarchical modules, resulting in additional performance gains with fewer parameters while being less computational.

Apart from this, Weijian Deng et al. [75] compared the performance and efficiency of information distillation to residual learning in lightweight SR models. They developed RepRFN, a re-parameterization-based lightweight SR network that minimizes GPU memory utilization while increasing inference performance. The network uses a multi-scale feature fusion structure to successfully collect different features and high-frequency edges, while redundant modules are removed to reduce complexity without losing performance. A Fourier transform-based loss function is also used to improve the learning of picture frequency components, allowing for effective real-time deployment on edge devices.

3.4. Deep Learning-Based Models for Real-Time Super-Resolution

In recent years, we have seen swift progress in the field of real-time SR thanks to enhancements in computational power and the emergence of deep learning. The purpose of deep learning is to discern the characteristic distribution of data by learning a layered representation of fundamental features. In more detail, deep learning progressively fine-tunes the SR algorithm methodology via an array of learning tactics. These include aspects like deep network architecture, optimization tools, and the design of loss functions. Concurrently, it also addresses the inherent challenges associated with SR. In the last years, many deep learning-based real-time SR models were introduced and made significant progress. In this section, we will in detail observe the most famous deep learning models such as CNNs, recurrent neural networks (RNNs), generative adversarial networks (GANs), and sparse representation models.

3.4.1. Deep Convolutional Neural Network-Based Models for Real-Time Super-Resolution

Initiallially, in ref. [49], an SR convolutional neural network (SRCNN) was described that uses end-to-end learning to map LR images to HR images. The SRCNN re-interprets existing sparse-coding-based SR approaches in the context of a deep convolutional neural network. Unlike traditional sparse coding techniques, which process each component independently, the SRCNN optimizes all layers simultaneously. The network is also light and quick, making it ideal for real-time web applications. Despite its end-to-end learning capabilities, the SRCNN's performance is limited by its shallow convolutional design and tiny receptive field. Following this [76], a very deep convolutional network (VDSR) for single-image SR (SISR) was built using the VGG-net [77]. The main idea underlying the VDSR is to improve accuracy by dramatically expanding the network depth. The VDSR design has 20 weight layers, which are built with a deep structure of recursively cascaded tiny filters. To expedite convergence, the network learns residuals at extraordinarily high rates—up to 10,000 times faster than those utilized in the SRCNN [49]. Additionally, gradient clipping is used to preserve training stability. The VDSR model does, however, have a restriction in that it is based on a fixed-size receptive field. In addition [78], a fast SR convolutional neural network (FSRCNN) has been introduced to speed up the SRCNN. The FSRCNN redesigns the SRCNN in three ways: it adds a deconvolution layer at the end to reduce computation by directly upsampling images; it uses a mapping layer with smaller filters after shrinking the input features to improve efficiency; and it optimizes parameters for real-time CPU performance. These enhancements enable the FSRCNN to outperform SRCNN in terms of processing speed and image restoration quality. However, these methods fail to produce better image quality, and the resulting SR images often appear blurred or exhibit edge halos. To address these issues, Zhang et al. [55] introduced an adaptive importance learning scheme model (VDSR-f22+ILT) proposed as an addition to the VDSR model, which incorporates adaptive importance learning (AIL) for SISR. During training, the model constantly adjusts the relevance of image pixels based on the loss function. A custom built significance penalty function was utilized to gradually raise the emphasis on individual pixels by solving a convex optimization problem. Training begins with simpler pixels and progresses to more complex ones, allowing the network to develop a robust initial capability. The primary novelty of the VDSR-f22+ILT model is its new training strategy, which incorporates significance learning into a joint optimization process during the training phase.

Recently, ref. [79] presented a novel residual local feature network (RLFN) for SISR. The RLFN learns residual local features using three convolutional layers and a reduced feature aggregation algorithm to find a compromise between performance and processing time. Furthermore, a contrastive loss function is used based on the idea that picking intermediate characteristics from the feature extractor might improve model performance. To improve the performance even more, a multi-stage warm-start training strategy is used with pretrained weights from previous stages. These achievements allowed the RLFN to win first place in the NTIRE 2022 efficient SR competition.

3.4.2. Recurrent Neural Networks (RNNs)-Based Models for Real-Time Super-Resolution

Qian Ning et al. [80] present an explainable technique to SISR based on RNNs known as the model-guided deep unfolding network (MoG-DUN). To overcome the coherence barrier, they employed a well-known image prior, the non-local auto-regressive model, to guide their deep neural network. This technique embeds deep denoising and non-local regularization as trainable modules in a deep learning framework, transforming the iterative model-based SISR process into a three-stage sequence of interconnected modules: denoising, non-local AR, and reconstruction. These modules use sophisticated approaches such as dense/skip connections and fast non-local implementations. MoG-DUN is explainable, highly accurate with less aliasing artifacts, computationally economical with fewer parameters, and capable of managing a wide range of degradations while deploying in real time.

3.4.3. Generative Adversarial Networks (GANs) Based Models for Real-Time Super-Resolution

Dazhao Zhou et al. [81] present a new GAN-based model for SR image reconstructions. Their method is based on Residual-in-Residual Dense Blocks (RRDBs) as the basic components, and they employed high-order degradation to simulate real-world image degradation. Furthermore, they introduced an innovative attention mechanism known as the All-Attention Mechanism (AAM), which effectively recognizes and uses crucial visual elements, expanding the number of pixels exploited. This strategy significantly increased the quality of the re-built images. In addition, Vibhu Bhatia et al. [82] proposed an approach utilizing an existing deep learning approach to accomplish real-time SR microscopy using a conventional GPU. They suggest a tiling approach, which utilizes GPU parallelism to speed training networks. They also offer simple changes to the design of the SRGAN's generator and discriminator. The researchers then assessed the model's result quality and runtime, proving that it is suitable for implementation across a variety of sectors, including low-end benchtop and mobile microscopes.

Discussion and Limitations

The rapid growth of deep learning-based models for real-time SR has sparked heated debate, notably over the trade-offs between model complexity, computing efficiency, and image quality. The emergence of models such as the RLFN, which relies on residual local features and decreased feature aggregation, has demonstrated a promising balance between performance and processing time, as proven by its first place win in the NTIRE 2022 efficient SR competition. However, this result raises the question as to whether such efficient models may sacrifice generalizability and robustness when dealing with more complicated and diverse image datasets.

In parallel, RNN techniques like MoG-DUN prioritize explainability and the incorporation of non-local auto-regressive models, which improve accuracy and eliminate artifacts in SISR. Despite their accuracy and capacity to manage various degradations, this raises concerns regarding the feasibility of using such advanced approaches in settings that demand quick processing. GANs provide a new dimension to the discussion, particularly with the introduction of models that include RRDBs and sophisticated attention mechanisms such as the AAM. These models produce excellent image quality by efficiently exploiting a larger number of pixels but at the expense of additional computational overhead. Efforts to ease these demands through GPU parallelism and simpler network designs have made progress, but there are still hurdles in developing models that can retain this high quality while still being practical for real-time application on lower-end devices and in mobile environments.

The continuing debate in the area centers on how to effectively manage these trade-offs in order to construct SR models that are not only efficient and accurate but that are also effective and responsive to a wide range of real-world scenarios. Achieving this balance is crucial for the widespread use of SR technologies in a variety of industries, including those with limited computing resources.

3.5. Quantization-Based Real-Time Super Resolution

We are currently observing a rapid expansion and pervasive implementation of deep neural networks across a multitude of disciplines. However, these DNN models, particularly CNNs, typically comprise an enormous number of parameters and demand high computational resources, tethering them closely to high-end hardware. This poses a significant hindrance to their broader applications, such as deployment on mobile devices.

3.5.1. Quantized CNN-Based Approaches

Mustafa Ayazoglu [17] presents XLSR, a unique, hardware-aware model built for real-time SR with an emphasis on extreme lightweight and robustness to quantization—specifically with the Synaptics Dolphin NPU in mind. The root modules established in the DeepRoots framework for image categorization served as inspiration for the model's fundamental design. The author successfully customized these root modules to tackle the SISR issue. To improve the model's resilience to uint8 quantization, they added a Clipped ReLU function to the network's last layer, resulting in an ideal combination of image reconstruction quality and processing time. Interestingly, despite having 30 times fewer parameters than the VDSR model, XLSR surpassed VDSR on the Div2K validation dataset. In [83], the authors suggest a novel approach for SR networks known as Content-Aware Dynamic Quantization (CADyQ), which aims to reduce large bits while minimizing accuracy loss. CADyQ dynamically assigns optimum bit widths to different local areas and network layers based on the particular information in the input image. A trainable bit selection module is presented to identify the optimal bit width and quantization level for each layer and local picture patch. This module is based on quantization sensitivity, which is computed by combining the average magnitude of the patch's image gradient alongside the standard deviation of the layer's input features.

Furthermore, Ziwei Luo et al. [45] introduce a basic convolutional network called NCNet, which includes a quick nearest-convolution module. This architecture is NPU-friendly, allowing for dependable real-time SR. Their technique uses the closest convolution, which provides the same performance as nearest upsampling but is significantly quicker and more suited to Android NNAPI. The model is easily deployable on mobile devices that support 8-bit quantization and is completely interoperable with all popular mobile AI accelerators. They also conducted extensive experiments on mobile devices using various tensor operations to demonstrate the effectiveness of their network design in real time.

Discussion and Limitations

The fast development of DNNs, particularly CNNs, has resulted in substantial advances in a variety of domains. However, the computational needs and huge parameter values of these models have restricted their use on resource-constrained devices such as mobile phones. To overcome this, numerous approaches have arisen to make CNNs more efficient while maintaining performance. The XLSR model, for example, has been created with hardware awareness in mind, allowing it to be used on specialized NPUs such as the Synaptic Dolphin. It includes a Clipped ReLU function, which improves resilience to quantization and achieves excellent performance with many fewer parameters. Similarly, CADyQ introduces dynamic quantization that optimally adjusts bit widths across network layers and local regions based on image content, further reducing computational load while preserving accuracy. NCNet presents an NPU-friendly architecture tailored for real-time SR on mobile devices. Despite these advancements, limitations remain. The dependence on specific hardware accelerators like NPUs may restrict wider deployment across different platforms. Moreover, dynamic quantization techniques, while effective, add complexity to the training process and may lead to performance trade-offs, especially in highly detailed tasks where extreme quantization can degrade accuracy.

4. Implementation and Analysis of Real-Time SR Algorithms in IoT Context

4.1. Implementation on IoT Hardware

Deploying SR algorithms on IoT-specific hardware presents unique challenges and opportunities. To gain practical insights into real-world applications, we explored approaches conducted on microcontrollers and low-end FPGAs.

For microcontroller implementations, processors like the ARM Cortex-M family offer a balance between power efficiency and computational performance. These processors are well suited for lightweight SR models, such as simplified versions of FSRCNN or ESPCN, which have shown potential in resource-constrained environments.

FPGA-based solutions have emerged as a promising alternative for deploying DNN-based SR algorithms at the edge. For instance, Liu et al. [84] proposed a lightweight version of the FSRCNN for FPGAs by replacing the deconvolutional layer with a convolutional layer, significantly reducing the memory footprint. However, this compression resulted in some performance degradation compared to the original FSRCNN. Additionally, NeuriCam [85] introduced a dual-mode IoT camera system that integrates a low-power (1.1 mW) grayscale camera with a duty-cycled high-power (100 mW) color camera. This system employs a neural network decoder on a nearby gateway to reconstruct HR color video from the LR grayscale stream supplemented by periodic HR keyframes.

4.2. Comparative Analysis

To evaluate the applicability, we present a comparison of real-time SR implementations in IoT systems using key metrics presented in Table 2.

Table 2. Comparison of SR methods on Raspberry Pi 3 for the scale factor of x4.

SR Method	Latency (ms)	Energy (mW)	PSNR	SSIM
Bicubic Interpolation	50	400	24.3	0.821
SRCNN	120	750	25.6	0.833
ESPCN	80	750	26.9	0.852
Liu et al. [84]	95	700	25.6	0.869
NeuriCam [85]	200	850	26.5	0.872

This comparison highlights the trade-offs between performance metrics across different SR methods implemented on a common IoT platform. Here, the ESPCN offers a better balance, surpassing the SRCNN in image quality (PSNR: 26.9, SSIM: 0.852) while consuming the same amount of energy (750 mW) and exhibiting shorter latency. The approach from Liu et al. adopted a balanced approach, achieving an improved SSIM (0.869) with 95 ms latency and 700 mW energy usage while maintaining a PSNR of 25.6. NeuriCam performed better than the other approaches (PSNR: 26.5, SSIM: 0.872) albeit with the highest latency (200 ms) and energy usage (850 mW).

4.3. Challenges in Hardware Deployment

Deploying SR algorithms on IoT devices poses numerous obstacles. One significant challenge is the low computing capability of IoT devices, which sometimes lack the processing capabilities needed for complicated SR algorithms. This constraint involves the creation of lightweight models or the use of optimization techniques like pruning and quantization. Another issue is memory limits, since many IoT devices have limited RAM and ROM. This limits the size and complexity of deployed SR models. For example, creating S-Boxes with combinational logic rather than look-up tables can drastically reduce memory needs [86].

Furthermore, energy efficiency is crucial for battery-powered IoT devices, which necessitate solutions with low power consumption. For example, NeuriCam's [85] dual-camera hardware prototype produces 640×480 color wireless video at 15 fps while consuming just 46 mA, demonstrating the promise for energy-efficient SR in IoT applications.

Subsequently, hardware heterogeneity in IoT networks [87] is a significant barrier. These ecosystems frequently comprise disparate hardware platforms with differing capabilities, making it challenging to provide universally compatible SR solutions. Addressing these problems is critical to the effective implementation of SR algorithms on IoT devices.

5. Future Roadmap

Recent advances in SISR technology are significantly improving image quality. Nonetheless, there is still room for growth in a multitude of sectors. Deep learning models are likely to improve in the future, particularly in terms of efficiency and real-time inference. The most anticipated future directions include advancements in model compression techniques, the exploration of adaptive and dynamic approaches, and the development of domain-specific real-time SR. Figure 7 shows the real-time SR technique improvements through various means expected to see in the future.

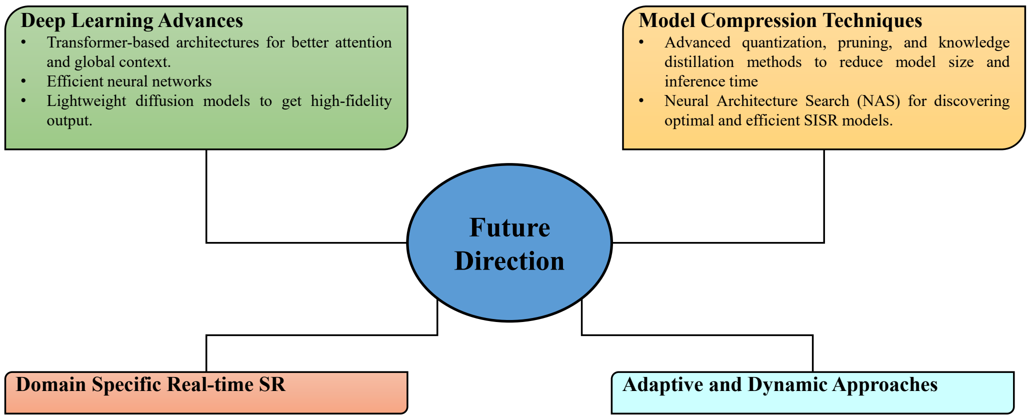

Figure 7. Future improvements in SR.

Advancements in deep learning models indicate that deep learning will continue to drive breakthroughs in SR research. Deep learning is expected to further enhance SR by developing models that explore sophisticated architectures, such as transformer-based lightweight structures, which represent a new frontier by employing enhanced attention mechanisms to improve both global context and local details while maintaining a balance of feature richness and network complexity. This innovation is critical for allowing SR in resource-constrained situations like edge devices and real-time applications. Simultaneously, effective neural networks can be tuned using techniques such as parameter pruning, quantization, and knowledge distillation, which reduce model complexity while retaining high output quality. In addition to these developments, lightweight diffusion models, which excel at iterative picture refining, have shown outstanding potential for SR applications by combining precision and efficiency. These advances are redefining the landscape of SR, assuring high-quality performance while tackling issues in real-time and resource-constrained applications.

In addition, model compression techniques are one of the key strategies for real-time SR, which aims to provide high-quality results in a timely way by utilizing advanced quantization, pruning, and knowledge distillation approaches to reduce model size and inference time without compromising the quality of output images. Quantization involves decreasing the accuracy of the model's weights and activations, which reduces memory and processing needs. Pruning gradually eliminates less-essential weights from the model, substantially simplifying its design while preserving performance. Knowledge distillation involves transferring knowledge from a bigger, more complicated model (the instructor) to a smaller, more efficient one (the student), allowing the smaller model to reach comparable levels of accuracy. These approaches are critical for ensuring that real-time SR systems can run quickly while maintaining image quality, making them useful for SISR. Apart from this, the Neural Architecture Search (NAS) for discovering optimal and efficient SISR models is another future direction that falls under model compression techniques.

The NAS automates the process of constructing neural networks by looking for the most efficient architecture that meets the unique needs of real-time SR activities. Using the NAS, researchers may uncover optimum designs that balance model complexity, speed, and accuracy. This enables the creation of unique models that are extremely efficient in terms of both computational cost and performance. Future research will most likely focus on merging these model compression approaches with NAS to develop even more optimal models for specific use cases, such as improving LR pictures in medical imaging, increasing satellite imagery quality, or recovering ancient photographs and films. Furthermore, there is a rising interest in adaptive and dynamic algorithms that may modify the amount of compression or model complexity based on the context or available computational resources, broadening the possible applications for real-time SR techniques.

Apart from these approaches, domain-specific real-time SR as well as adaptive and dynamic approaches for SR, are emerging trends in real-time SR. Domain-specific real-time SR focuses on applying SR methods to specific applications, such as medical imaging, satellite photography, or video streaming, where unique limits and requirements necessitate customized models. These tailored models can improve accuracy and efficiency by taking use of the target domain's unique properties. Meanwhile, adaptive and dynamic methods to SR entail creating models that can modify their complexity and processing capacity in real time based on available computing resources or task-specific requirements. These strategies are intended to improve performance in a variety of contexts, ranging from high-performance servers to low-power constrain devices, while maintaining constant quality and speed. These changes, taken together, imply a move toward more adaptable, efficient, and application-specific solutions in real-time SR.

6. Conclusions

This survey provides an in-depth examination of the majority of studies on real-time SR published up to 2024, presenting a structured and comprehensive taxonomy to classify and analyze these approaches. It delves into the intricacies of various methods, offering a nuanced discussion of their strengths and limitations as highlighted in the reviewed publications. Furthermore, it identifies and outlines the most frequently adopted strategies for addressing specific challenges in real-time SR, enabling a detailed understanding of the problem-solving methodologies employed in this domain.

Over the past decade, research in real-time SR has made notable strides, driven by advancements in deep learning and computational techniques. Nevertheless, the number of deep learning-based SR methods that achieve real-time performance remains limited. This shortfall can largely be attributed to the significant computational demands of these models, which often struggle to balance rapid processing speeds with high-quality outputs, particularly on resource-constrained devices. The challenge is further compounded by the difficulty of designing architectures that are lightweight yet powerful enough to handle the intricate details required for SR tasks. These limitations underscore the need for novel strategies that combine computational efficiency with robust performance.

This review highlights several promising directions for future research. For instance, by integrating model compression techniques such as pruning, quantization, and knowledge distillation, hardware-aware NAS-based approcahes could yield lightweight, high-performance models optimized for real-time applications. Additionally, developing adaptive algorithms capable of dynamically adjusting model complexity based on available computational resources could significantly expand the applicability of real-time SR techniques. Domain-specific SR methods, tailored to the unique requirements of applications such as medical imaging, satellite imagery, or video streaming, also hold immense potential for further refinement and optimization.

By consolidating current knowledge and providing critical insights, this survey aims to serve as a foundational reference for researchers and developers seeking to enhance SR solutions. Whether the goal is improving runtime efficiency, achieving higher accuracy, or addressing application-specific challenges, this review offers practical guidance and a roadmap for advancing real-time SR. Ultimately, we hope this work inspires new innovations that will address existing gaps and unlock the full potential of super-resolution technology.

Funding: This research was supported by the Culture, Sports, and Tourism R&D Program through the Korea Creative Content Agency grant funded by the Ministry of Culture, Sports, and Tourism in 2023 (Project Name: Cultural Technology Specialist Training and Project for Metaverse Game, Project Number: RS-2023-00227648, Contribution Rate: 100%).

Data Availability Statement: The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Salvetti, F.; Mazzia, V.; Khaliq, A.; Chiaberge, M. Multi-image super resolution of remotely sensed images using residual attention deep neural networks. *Remote Sens.* **2020**, *12*, 2207. [CrossRef]
2. Gendy, G.; Mohammed, H.; Sabor, N.; He, G. A deep pyramid attention network for single image super-resolution. In Proceedings of the 2021 9th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC), Alexandria, Egypt, 13–14 December 2021; pp. 14–19.
3. Mardieva, S.; Ahmad, S.; Umirzakova, S.; Rasool, M.A.; Whangbo, T.K. Lightweight image super-resolution for IoT devices using deep residual feature distillation network. *Knowl. Based Syst.* **2024**, *285*, 111343. [CrossRef]
4. Rasool, M.A.; Jeong, W.; Ahmed, S.; Whangbo, T.K. Stellar SR: A Convolutional Local Feature Network for Lightweight Image Super-Resolution. *Korean Soc. Broad Eng.* **2024**, 298–301.
5. Chen, Y.; Liu, L.; Phonevilay, V.; Gu, K.; Xia, R.; Xie, J.; Zhang, Q.; Yang, K. Image super-resolution reconstruction based on feature map attention mechanism. *Appl. Intell.* **2021**, *51*, 4367–4380. [CrossRef]
6. Zeng, K.; Ding, S.; Jia, W. Single image super-resolution using a polymorphic parallel CNN. *Appl. Intell.* **2019**, *49*, 292–300. [CrossRef]
7. Jo, Y.; Oh, S.W.; Kang, J.; Kim, S.J. Deep video super-resolution network using dynamic upsampling filters without explicit motion compensation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 3224–3232.
8. Rasool, M.A.; Ahmed, S.; Sabina, U.; Whangbo, T.J. MRESR: Multi-agent Reinforcement learning for Efficient Super-Resolution. *Korean Soc. Broad Eng.* **2023**, 460–463.
9. Thawakar, O.; Patil, P.W.; Dudhane, A.; Murala, S.; Kulkarni, U. Image and video super resolution using recurrent generative adversarial network. In Proceedings of the 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Taipei, Taiwan, 18–21 September 2019; pp. 1–8.
10. Van Reeth, E.; Tham, I.W.; Tan, C.H.; Poh, C.L. Super-resolution in magnetic resonance imaging: A review. *Concepts Magn. Reson. Part* **2012**, *40*, 306–325. [CrossRef]
11. Rasool, M.A.; Ahmed, S.; Sabina, U.; Whangbo, T.K. KONet: Towards a Weighted Ensemble Learning Model for Knee Osteoporosis Classification. *IEEE Access* **2024**, *12*, 5731–5742. [CrossRef]
12. Smith, E.; Fujimoto, S.; Meger, D. Multi-view silhouette and depth decomposition for high resolution 3d object representation. *Adv. Neural Inf. Process. Syst.* **2018**, *31*.
13. Zhang, K.; Danelljan, M.; Li, Y.; Timofte, R.; Liu, J.; Tang, J.; Wu, G.; Zhu, Y.; He, X.; Xu, W.; et al. Aim 2020 challenge on efficient super-resolution: Methods and results. In Proceedings of the Computer Vision–ECCV 2020 Workshops, Glasgow, UK, 23–28 August 2020; Proceedings, Part III 16; Springer: Berlin/Heidelberg, Germany, 2020; pp. 5–40.
14. Liu, A.; Liu, Y.; Gu, J.; Qiao, Y.; Dong, C. Blind image super-resolution: A survey and beyond. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 5461–5480. [CrossRef] [PubMed]
15. Liu, J.; Tang, J.; Wu, G. Residual feature distillation network for lightweight image super-resolution. In Proceedings of the Computer Vision–ECCV 2020 Workshops, Glasgow, UK, 23–28 August 2020; Proceedings, Part III 16; Springer: Berlin/Heidelberg, Germany, 2020; pp. 41–55.

16. Wang, Y.; Wang, L.; Yang, J.; An, W.; Guo, Y. Flickr1024: A large-scale dataset for stereo image super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, Long Beach, CA, USA, 15–20 June 2019.
17. Ayazoglu, M. Extremely lightweight quantization robust real-time single-image super resolution for mobile devices. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 2472–2479.
18. Fang, J.; Lin, H.; Chen, X.; Zeng, K. A hybrid network of cnn and transformer for lightweight image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1103–1112.
19. Bulat, A.; Tzimiropoulos, G. Super-fan: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with gans. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018 ; pp. 109–117.
20. Bei, Y.; Damian, A.; Hu, S.; Menon, S.; Ravi, N.; Rudin, C. New techniques for preserving global structure and denoising with low information loss in single-image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018 pp. 874–881.
21. Zhang, K.; Zuo, W.; Zhang, L. Learning a single convolutional super-resolution network for multiple degradations. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018 pp. 3262–3271.
22. Behjati, P.; Rodriguez, P.; Mehri, A.; Hupont, I.; Tena, C.F.; Gonzalez, J. Overnet: Lightweight multi-scale super-resolution with overscaling network. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Virtual, 11–17 October 2021; pp. 2694–2703.
23. Fan, Y.; Shi, H.; Yu, J.; Liu, D.; Han, W.; Yu, H.; Wang, Z.; Wang, X.; Huang, T.S. Balanced two-stage residual networks for image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 161–168.
24. Tai, Y.; Yang, J.; Liu, X. Image super-resolution via deep recursive residual network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 3147–3155.
25. Wang, L.; Dong, X.; Wang, Y.; Ying, X.; Lin, Z.; An, W.; Guo, Y. Exploring sparsity in image super-resolution for efficient inference. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 4917–4926.
26. Chen, R.; Qu, Y.; Zeng, K.; Guo, J.; Li, C.; Xie, Y. Persistent memory residual network for single image super resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 809–816.
27. Wang, Y.; Perazzi, F.; McWilliams, B.; Sorkine-Hornung, A.; Sorkine-Hornung, O.; Schroers, C. A fully progressive approach to single-image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 864–873.
28. Anwar, S.; Khan, S.; Barnes, N. A deep journey into super-resolution: A survey. *Acm Comput. Surv.* **2020**, *53*, 1–34. [CrossRef]
29. Wang, X.; Yu, K.; Dong, C.; Loy, C.C. Recovering realistic texture in image super-resolution by deep spatial feature transform. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 606–615.
30. Timofte, R.; Rothe, R.; Van Gool, L. Seven ways to improve example-based single image super resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1865–1873.
31. Matsui, Y.; Ito, K.; Aramaki, Y.; Fujimoto, A.; Ogawa, T.; Yamasaki, T.; Aizawa, K. Sketch-based manga retrieval using manga109 dataset. *Multimed. Tools Appl.* **2017**, *76*, 21811–21838. [CrossRef]
32. Gendy, G.; He, G.; Sabor, N. Lightweight image super-resolution based on deep learning: State-of-the-art and future directions. *Inf. Fusion* **2023**, *94*, 284–310. [CrossRef]
33. Chen, H.; He, X.; Qing, L.; Wu, Y.; Ren, C.; Sheriff, R.E.; Zhu, C. Real-world single image super-resolution: A brief review. *Inf. Fusion* **2022**, *79*, 124–145. [CrossRef]
34. Martin, D.; Fowlkes, C.; Tal, D.; Malik, J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In Proceedings of the Eighth IEEE International Conference on Computer Vision, ICCV 2001, Vancouver, BC, Canada, 7–14 July 2001; Volume 2, pp. 416–423.
35. Yuan, Y.; Liu, S.; Zhang, J.; Zhang, Y.; Dong, C.; Lin, L. Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–23 June 2018; pp. 701–710.
36. Aly, H.A.; Dubois, E. Image up-sampling using total-variation regularization with a new observation model. *IEEE Trans. Image Process.* **2005**, *14*, 1647–1659. [CrossRef] [PubMed]
37. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.

38. Wang, Z.; Simoncelli, E.P.; Bovik, A.C. Multiscale structural similarity for image quality assessment. In Proceedings of the Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, USA, 9–12 November 2003; Vourne 2, pp. 1398–1402.
39. Song, D.; Xu, C.; Jia, X.; Chen, Y.; Xu, C.; Wang, Y. Efficient residual dense block search for image super-resolution. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 12007–12014.
40. Zong, Z.; Zha, L.; Jiang, J.; Liu, X. Asymmetric information distillation network for lightweight super resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1249–1258.
41. Li, K.; Yang, S.; Dong, R.; Wang, X.; Huang, J. Survey of single image super-resolution reconstruction. *Iet Image Process.* **2020**, *14*, 2273–2290. [CrossRef]
42. Shi, W.; Caballero, J.; Huszár, F.; Totz, J.; Aitken, A.P.; Bishop, R.; Rueckert, D.; Wang, Z. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1874–1883.
43. Tipping, M.; Bishop, C. Bayesian image super-resolution. *Adv. Neural Inf. Process. Syst.* **2002**, *15*.
44. Yao, X.; Wu, Q.; Zhang, P.; Bao, F. Adaptive rational fractal interpolation function for image super-resolution via local fractal analysis. *Image Vis. Comput.* **2019**, *82*, 39–49. [CrossRef]
45. Luo, Z.; Li, Y.; Yu, L.; Wu, Q.; Wen, Z.; Fan, H.; Liu, S. Fast nearest convolution for real-time efficient image super-resolution. In Proceedings of the Computer Vision–ECCV 2022 Workshops, Tel Aviv, Israel, 23–27 October 2022; Proceedings, Part II; Springer: Berlin/Heidelberg, Germany, 2023; pp. 561–572.
46. Learner, T.A. Image Processing—Nearest Neighbour Interpolation. Available online: https://theailearner.com/2018/12/29/image-processing-nearest-neighbour-interpolation/ (accessed on 9 May 2023).
47. Khaledyan, D.; Amirany, A.; Jafari, K.; Moaiyeri, M.H.; Khuzani, A.Z.; Mashhadi, N. Low-cost implementation of bilinear and bicubic image interpolation for real-time image super-resolution. In Proceedings of the 2020 IEEE Global Humanitarian Technology Conference (GHTC), Seattle, WA, USA, 29 October–1 November 2020; pp. 1–5.
48. Cho, H.M.; Choi, K. Super-Resolution based Video Coding Scheme. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1778–1780.
49. Dong, C.; Loy, C.C.; He, K.; Tang, X. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *38*, 295–307. [CrossRef]
50. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4681–4690.
51. Ganapati, V. Optimal physical preprocessing for example-based super-resolution (Conference Presentation). In *Optical Data Science II*; SPIE: Bellingham, WA, USA, 2019; Volume 10937, p. 109370D.
52. Lim, B.; Son, S.; Kim, H.; Nah, S.; Lee, K.M. Enhanced deep residual networks for single image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 136–144.
53. Lai, W.S.; Huang, J.B.; Ahuja, N.; Yang, M.H. Deep laplacian pyramid networks for fast and accurate super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 624–632.
54. Zhang, K.; Li, D.; Luo, W.; Ren, W.; Stenger, B.; Liu, W.; Li, H.; Yang, M.H. Benchmarking ultra-high-definition image super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Virtual, 11–17 October 2021; pp. 14769–14778.
55. Zhang, L.; Wang, P.; Shen, C.; Liu, L.; Wei, W.; Zhang, Y.; Van Den Hengel, A. Adaptive importance learning for improving lightweight image super-resolution network. *Int. J. Comput. Vis.* **2020**, *128*, 479–499. [CrossRef]
56. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1026–1034.
57. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef]
58. Sheikh, H.R.; Bovik, A.C.; De Veciana, G. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Trans. Image Process.* **2005**, *14*, 2117–2128. [CrossRef] [PubMed]
59. Zhang, R.; Isola, P.; Efros, A.A.; Shechtman, E.; Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 586–595.
60. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a "completely blind" image quality analyzer. *IEEE Signal Process. Lett.* **2012**, *20*, 209–212. [CrossRef]
61. Yang, C.Y.; Ma, C.; Yang, M.H. Single-image super-resolution: A benchmark. In Proceedings of the Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, 6–12 September 2014; Proceedings, Part IV 13; Springer: Berlin/Heidelberg, Germany, 2014; pp. 372–386.

62. Red Hat. What Is a Latency-Sensitive Application? 2024. Available online: https://www.redhat.com/en/topics/edge-computing/latency-sensitive-applications#what-is-a-latency-sensitive-application (accessed on 15 August 2024).
63. Hui, Z.; Wang, X.; Gao, X. Fast and accurate single image super-resolution via information distillation network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 723–731.
64. Moon, H.C.; Kim, J.G.; Jeong, J.; Kim, S. Feature-domain adaptive contrastive distillation for efficient single image super-resolution. *IEEE Access* **2023**, *11*, 131885–131896. [CrossRef]
65. Hui, Z.; Gao, X.; Yang, Y.; Wang, X. Lightweight image super-resolution with information multi-distillation network. In Proceedings of the 27th Acm International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 2024–2032.
66. Yang, H.; Jeon, G.; Liu, K.; Liu, Y.; Yang, X. Feature similarity rank-based information distillation network for lightweight image superresolution. *Knowl.-Based Syst.* **2023**, *266*, 110437. [CrossRef]
67. Yoon, M.; Lee, S.; Song, B.C. TAKDSR: Teacher Assistant Knowledge Distillation Framework for Graphics Image Super-Resolution. *IEEE Access* **2023**, *11*, 112015–112026. [CrossRef]
68. Sun, L.; Liang, J.; Liu, S.; Yong, H.; Zhang, L. Perception-Distortion Balanced Super-Resolution: A Multi-Objective Optimization Perspective. *IEEE Trans. Image Process.* **2024**, *33*, 4444–4458. [CrossRef] [PubMed]
69. Koshelev, I.; Somov, A. Learned Gradient Descent Performance in Bicubic Super-Resolution Task. In Proceedings of the IECON 2021–47th Annual Conference of the IEEE Industrial Electronics Society, Toronto, ON, Canada, 13–16 October 2021; pp. 1–6.
70. Li, W.; Li, M.; Zuo, L.; Chen, H.; Wu, Y.; Zhuo, Z. A computationally efficient airborne forward-looking super-resolution imaging method based on sparse Bayesian learning. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–13. [CrossRef]
71. Song, H.; Yang, Y. Uncertainty quantification in super-resolution guided wave array imaging using a variational Bayesian deep learning approach. *Ndt Int.* **2023**, *133*, 102753. [CrossRef]
72. Sarangi, P.; Hücümenoğlu, M.C.; Rajamäki, R.; Pal, P. Super-resolution with sparse arrays: A nonasymptotic analysis of spatiotemporal trade-offs. *IEEE Trans. Signal Process.* **2023**, *71*, 4288–4302. [CrossRef]
73. Ying, W.y.; Dong, T.y.; Fan, J. Edge-priority-extraction network using re-parameterization for real-time super-resolution. *Vis. Comput.* **2023**, *40*, 7623–7636. [CrossRef]
74. Wang, Y.; Su, T.; Li, Y.; Cao, J.; Wang, G.; Liu, X. DDistill-SR: Reparameterized dynamic distillation network for lightweight image super-resolution. *IEEE Trans. Multimed.* **2022**, *25*, 7222–7234. [CrossRef]
75. Deng, W.; Yuan, H.; Deng, L.; Lu, Z. Reparameterized residual feature network for lightweight image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 1712–1721.
76. Kim, J.; Lee, J.K.; Lee, K.M. Accurate image super-resolution using very deep convolutional networks. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1646–1654.
77. Simonyan, K. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
78. Dong, C.; Loy, C.C.; Tang, X. Accelerating the super-resolution convolutional neural network. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Proceedings, Part II 14; Springer: Berlin/Heidelberg, Germany, 2016; pp. 391–407.
79. Kong, F.; Li, M.; Liu, S.; Liu, D.; He, J.; Bai, Y.; Chen, F.; Fu, L. Residual local feature network for efficient super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 766–776.
80. Ning, Q.; Dong, W.; Shi, G.; Li, L.; Li, X. Accurate and lightweight image super-resolution with model-guided deep unfolding network. *IEEE J. Sel. Top. Signal Process.* **2020**, *15*, 240–252. [CrossRef]
81. Zhou, D.; Sun, S.; Li, M.; Liu, X.; Mi, S. GAN-based super-resolution for confocal superficial eyelid imaging: Real-time, domain generalization, and noise robustness. *Biomed. Signal Process. Control* **2024**, *89*, 105668. [CrossRef]
82. Bhatia, V.; Kumar, Y. Attaining real-time super-resolution for microscopic images using gan. *arXiv* **2020**, arXiv:2010.04634.
83. Hong, C.; Baik, S.; Kim, H.; Nah, S.; Lee, K.M. Cadyq: Content-aware dynamic quantization for image super-resolution. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 367–383.
84. Liu, H.; Qian, Y.; Liang, Y.; Zhang, B.; Liu, Z.; He, T.; Zhao, W.; Lu, J.; Yu, B. A High-Performance Accelerator for Real-Time Super-Resolution on Edge FPGAs. *Acm Trans. Des. Autom. Electron. Syst.* **2024**, *29*, 1–25. [CrossRef]
85. Veluri, B.; Pernu, C.; Saffari, A.; Smith, J.; Taylor, M.; Gollakota, S. NeuriCam: Key-Frame Video Super-Resolution and Colorization for IoT Cameras. In Proceedings of the 29th Annual International Conference on Mobile Computing and Networking, Madrid, Spain, 2–6 October 2023; pp. 1–17.

86. Almutairi, R.; Bergami, G.; Morgan, G. Advancements and Challenges in IoT Simulators: A Comprehensive Review. *Sensors* **2024**, *24*, 1511. [CrossRef] [PubMed]
87. Bouchemal, N.; Kallel, S.; Bouchemal, N. A survey: WSN heterogeneous architecture platform for IoT. In Proceedings of the Machine Learning for Networking: First International Conference, MLN 2018, Paris, France, 27–29 November 2018; Revised Selected Papers 1; Springer: Berlin/Heidelberg, Germany, 2019; pp. 321–332.